

МНОГОКОМПОНЕНТНЫЙ АНАЛИЗ ПРОГНОЗИРОВАНИЯ МЕТЕОРОЛОГИЧЕСКИХ РЯДОВ С ИСПОЛЬЗОВАНИЕМ РЕКУРРЕНТНЫХ НЕЙРОННЫХ СЕТЕЙ

Е.В. Унжакова, И.А. Ботыгин
Томский политехнический университет
E-mail: katerina.unz@gmail.com

Введение

Задача прогнозирования будущих значений весьма актуальна в наши дни и её решение играет важнейшую роль в процессах как стратегического планирования, так и оперативного управления в различных сферах науки, техники, медицины, экономики, политики и т.д. [1-3].

Одна из моделей для прогнозирования временных рядов, получившая широкое распространение – это модель на искусственных нейронных сетях (ИНС). Способность приложений на базе искусственных нейронных сетей к обобщению и выделению скрытых зависимостей между входными и выходными данными предоставляет уникальные возможности и для задач прогнозирования [4].

В настоящей работе приведены результаты сравнительного многокомпонентного анализа прогнозирования метеорологических рядов с использованием рекуррентных ИНС.

Сравнение результатов при однокомпонентном и многокомпонентном прогнозировании

Эксперимент был проведен на облачном сервисе Google Cloud Platform [5-7]. Это весьма эффективно и удобно в том плане, что обучение может проводиться полностью независимо от пользователя и вычислительных возможностей его ПК. Кроме того, для доступа к облаку было достаточно мобильного устройства для проверки результатов обучения.

В облаке была создана виртуальная вычислительная машина, на которую были установлены необходимые компоненты.

Использовался набор метеорологических данных с сайта Всероссийского научно-исследовательского института гидрометеорологической информации. Для эксперимента были выбраны данные по станции Томск за последние 10 лет (до 2018 г.) по 24 признакам [8].

Наблюдения в выборке каждые 3 часа – это 8 наблюдений за 24 часа. Прогнозирование осуществлялось на период в 7 дней. В соответствии с выбранным сроком прогнозирования (на 7 следующих дней) сделано смещение по времени в данных на 56 шагов ($7 \times 8 = 56$).

ИНС обучалась 3 раза. В первом случае прогнозировались значения по трём компонентам: температура воздуха по сухому термометру, относительная влажность воздуха и атмосферное давление на уровне станции. Во втором – по двум

компонентам: температура воздуха по сухому термометру и относительная влажность воздуха. В третий раз – однокомпонентное прогнозирование температуры воздуха по сухому термометру.

В целях ускорения процесса обучения ИНС из полной базы учебных данных были созданы группы более коротких последовательностей данных для обучения.

Для удобной визуализации, рассмотрим на графиках результаты прогнозирования значений температуры воздуха по сухому термометру, при обучении на тестовых данных, которых нейросеть не видела в процессе обучения (рис. 1-3).

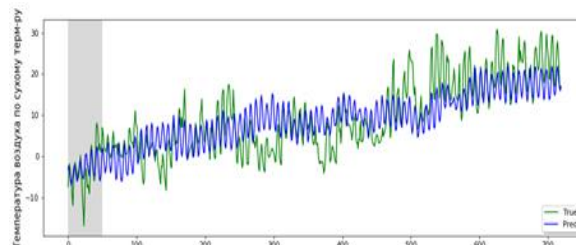


Рис. 1. Результаты обучения ИНС по трём признакам при прогнозировании значений температуры воздуха по сухому термометру

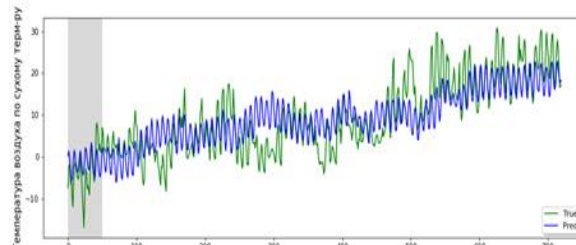


Рис. 2. Результаты обучения ИНС по двум признакам при прогнозировании значений температуры воздуха по сухому термометру

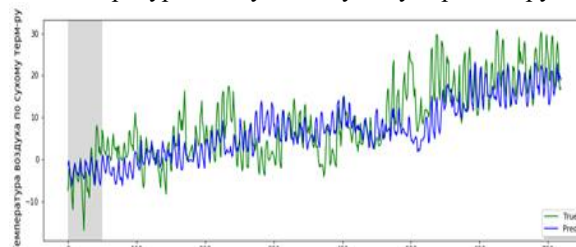


Рис. 3. Результаты обучения ИНС при однокомпонентном прогнозировании температуры воздуха по сухому термометру

На графиках представлены результаты за 3 месяца (720 измерений). По оси X указано время (единица измерения времени – 3 часа). По оси Y

указаны значения от минимального до максимального значения по конкретному предсказанному признаку – температура воздуха по сухому термометру в градусах Цельсия.

Обозначение «True.» – настоящие результаты (зелёный), «Pred.» – предсказанные ИНС (синий). Серым цветом выделен период разогрева.

Из графиков видно, что ИНС достаточно точно предсказывает значения при многокомпонентном прогнозировании (графики визуально не изменились), однако при однокомпонентном прогнозировании точность снизилась.

При обучении модели были использованы функции потерь для оценки качества прогнозирования:

- средняя квадратичная ошибка (Mean Squared Error, MSE). Она диагностирует, как близко вывод модели соответствует настоящим выходным сигналам [9],
- метрика ассигасу, добавленная при компиляции модели, которая показывает точность предсказания на тестовом наборе данных [10].

Представим результаты производительности прогнозирования модели ИНС на тестовом наборе на рисунке 4.

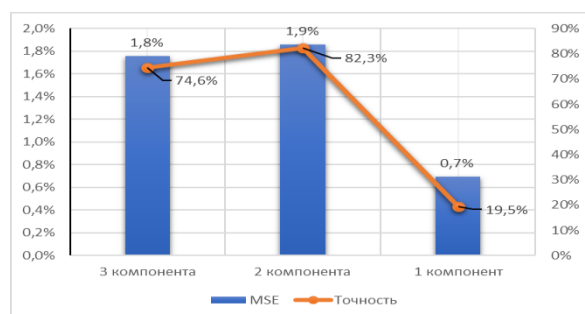


Рис. 4. Зависимость показателей MSE и точности предсказания от количества компонентов предсказания

По итогам при многокомпонентном обучении MSE – 1,8% и точность предсказания – 74,6% с тремя прогнозируемыми признаками, 1,9% и 82,3% соответственно с двумя признаками. Что является хорошим результатом. Во время предсказания одного признака нейросеть показала MSE равной 0,7%, однако метрика ассигасу равнялась нулю.

Тип созданной модели подходит для предсказания временных рядов с задачей классификации входных значений, следовательно, применение метрики ассигасу становится неуместным. Поэтому была создана функция для оценки точности предсказания ($|\text{Pred.} - \text{True.}|$), которая показала очень низкий процент – 19,5%.

Заключение

Исследование показало определённую эффективность при многокомпонентном прогнозировании для решения задач

прогнозирования метеорологических временных рядов с помощью рекуррентных нейронных сетей.

Варьируя количество эпох, шагов за эпоху и длину последовательностей можно добиться значительного улучшения качества предсказания и точности в процентном отношении. При выборе одного признака прогнозирования точность предсказания созданной модели значительно снижается.

Список использованных источников

1. Роль и сущность прогнозирования // [Электронный ресурс]. – URL: <http://finans-study.ru/rol-i-sushhnost-prognostirovaniya.html> (дата обращения: 14.11.2019).
2. Обзор методов прогнозирования // [Электронный ресурс]. – URL: <https://ivan-shamaev.ru/overview-forecast-methods/#i-14> (дата обращения: 15.11.2019).
3. Комплексное прогнозирование научно-технологического развития // [Электронный ресурс]. – URL: <https://cyberleninka.ru/article/n/kompleksnoe-prognostirovanie-nauchno-tehnologicheskogo-razvitiya-opyt-i-uroki> (дата обращения: 15.11.2019).
4. Антонио, Д. Библиотека Keras – инструмент глубокого обучения. Реализация нейронных сетей с помощью библиотек Theano и TensorFlow / Д. Антонио, П. Суджит; перевод с английского А.А. Слинкин. — Москва: ДМК Пресс, 2018. — 294 с.
5. Топ10+ зарубежных облачных провайдеров // [Электронный ресурс]. – URL: <http://cloudzone.ru/articles/review/2.html> (дата обращения: 17.11.2019).
6. Платформа облачных вычислений – Google Cloud Platform // [Электронный ресурс]. – URL: <https://cloud.google.com/> (дата обращения: 17.11.2019).
7. Тренируем нейронную сеть, написанную на TensorFlow в облаке // [Электронный ресурс]. – URL: <https://habr.com/ru/post/318922> (дата обращения: 17.11.2019).
8. Специализированные массивы для климатических исследований // [Электронный ресурс]. – URL: <http://aisori-m.meteo.ru/waisori/index.xhtml?idata> (дата обращения: 17.11.2019).
9. Жерон, Орельен. Прикладное машинное обучение с помощью Scikit-Learn и TensorFlow. Пер. с англ. — СПб.: ООО «Альфа-книга»: 2018. — 688 с.
10. Usage of metrics – Keras Documentation // [Электронный ресурс]. – URL: <https://keras.io/metrics/#accuracy> (дата обращения: 25.11.2019).